Intelligent Agents

Fourier Analysis II: Social Theory and Proof of Arrow‘s Theorem

Özgür L. Özçep

Universität zu Lübeck

Institut für Informationssysteme

Todays and next weeks lecture based on

- Lecture notes „Fourier Analysis of Boolean Functions, Winter term 16/17“
  M. Schweighofer
  http://www.math.uni-konstanz.de/~schweigh/

- Ryan O‘Donnell: Fourier Analyis of Boolean Functions., CUP 2014.
  Free PDF available at
  https://arxiv.org/pdf/2105.10386.pdf

- Talk of Ronald de Wolf: „Fourier analysis of Boolean functions: Some beautiful examples“ available at
  https://nvti.nl/slides/deWolf.pdf

SOCIAL CHOICE IN FOURIER ANAYLYIS

Motivation

- Social theory can be elegantly treated with Fourier Analysis
- The main aim of this lecture is:
  - sketch the basic ideas on Fouries analysis treatment of social theory
  - and as a highlight demonstrate G. Kalai‘s Fourier-theoretic proof of Arrow‘s theorem (Kalai 02)
- As a side effect (as before) we will see the power of the probabilistic method. (Nalon/Spencer 04)

Social Choice

$f: \{1,-1\}^n \ \rightarrow \{1,-1\}$ as voting rule with 2 candidates and n voters

Definition

- $maj_n(x) = \text{sgn}(x_1 + \cdots + x_n)$ (Majority function)
  (for n even $f(0)$ assigned arbitrarily).
- $f(x) = \text{sgn}(a_1 x_1 + \cdots + a_2 x_n)$ (weighted majority/
  for some $a \ \in \{1,-1\}^n$ linear threshold)
- $AND_n(x) = +1$ unless all $x_i = -1$ (AND function)
- $OR_n(x) \quad = -1$ unless all $x_i = 1$ (OR function)
- $\chi_i(x) = x_i$ (ith dictator function)
- $f(x) = g(x_{i_1}, \dots, x_{i_k})$ (k-junta)
  for some $g: \{1,-1\}^k \rightarrow \{1,-1\}$ and $\{i_1, \dots, i_k\} \in [n]$

Social Choice

$f: \{1,-1\}^n \to \{1,-1\}$ as voting rule with 2 candidates and n voters

Definition

- $maj_n^{\otimes d+1}(x^{(1)}, \dots, x^{(n)})$ (depth-d recursive majority)
  $= \quad mai_n\left(maj_n^{\otimes d}(x^{(1)}), \dots, maj_n^{\otimes d}(x^{(n)})\right)$
  (for $d \in \mathbb{N}_0$, n odd, and all $x^{(i)} \in \{1,-1\}^{n^d}$)

- $Tribes_{w,s}: \{1,-1\}^{ws} \to \{1,-1\}$ (tribes function)
  $Tribes_{w,s}(x^{(1)}, \dots, x^{(s)}) =$
  $OR_s\left(AND_w(x^{(1)}), \dots, AND_w(x^{(s)})\right)$
  (for $w, s \in \mathbb{N}_0, x^{(i)} \in \{1,-1\}^w$)

- Depth-2 recursive majority used in presidental elections in USA

Social Choice

$f: \{1, -1\}^n \rightarrow \{1, -1\}$ as voting rule with 2 candidates and n voters

Definition

$f: \{1, -1\}^n \rightarrow \mathbb{R}$ is called

- monotone iff $\mathrm{f}(\mathrm{x}) \leq f(y)$ for all $x, y \in \{1, , -1\}^n$ with $x_i \leq y_i$ for all $i \in [n]$
- odd iff $f(x) = -f(-x)$ for all $x \in \{1, -1\}^n$
- unanimous iff $f(1, \ldots, 1) = 1$ and $f(-1, \ldots, -1) = -1$
- symmetric iff $f\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = f(x)$ for all $x \in \{1, -1\}^n$ and all permutations $\sigma \in S_n$
- Transitive-symmetric iff for all i, j $\in [n]$ there is some $\sigma \in S_n$ such that $\sigma(i) = j$ and $f\left(x_{\sigma(1)}, \ldots, x_{\sigma(n)}\right) = f(x)$ for all $x \in \{1, -1\}^n$

Properties fulfilled by functions

Example

- For odd n maj_n has all properties and is the only monotone odd symmetric Boolen function on n bits
- maj_n (for odd n), AND_n, OR_n, and χ_i (for $i \in [n]$) are Boolean linear threshold functions
- AND_n, OR_nsatisfy all properties except oddness for $n \neq 1$ and unanimity for $n = 0$
- Dictator functions satisfy first three porperties but for $n \geq 2$ they do not satisfy the last two.

Properties fulfilled by functions

Example

- There are exactly $2n + 2$ 1-juntas on n bits (namely the n dictators, the n negated dictators and the two constant functions)

- For all $d \in \mathbb{N}_0$ and for odd n, $maj_n^{\otimes d}$ satisfies all properties except, in case of $n \geq 3$ and $d \geq 2$, symmetry

- For $w, s \in \mathbb{N}_{\geq 2}$, $\mathrm{T}ribes_{w,s}$ is monotone, not odd, unanimous, not symmetric but transitive symmetric.

INFLUENCES AND DERIVATIVES

i-th Influence

Definition

- For $x \in \{1,-1\}^n$, $i \in [n]$ and $b \in \{1,-1\}$ let
  - $x^{\oplus i} = (x_1, \dots, x_{i-1}, -x_i, x_{i+1}, \dots x_n)$
  - $x^{i \mapsto b} = (x_1, \dots, x_{i-1}, b, x_{i+1}, \dots x_n)$
- i is called pivotal for $f: \{1,-1\}^n \rightarrow \{1,-1\}$ on input x iff $f(x) \neq f(x^{\oplus i})$
- Influence of coordinate $i \in [n]$ on $f: \{1,-1\}^n \rightarrow \{1,-1\}$ is the probability that i is pivotal on a random input
  $Inf_i(f) = \Pr_{x \sim \{1,-1\}^n}(f(x) \neq f(x^{\oplus i}))$

Example

- $Inf_i(OR_n) = Inf_i(AND_n) = 2^{1-n}$
- $Inf_i(maj_n) = \binom{n-1}{\frac{n-1}{2}} 2^{1-n}$ (for odd n)

Definition

- The i-th derivative operator is defined by

  $D_i\colon\ \mathbb{R}^{\{1,-1\}^n} \to \mathbb{R}^{\{1,-1\}^n}, f \mapsto \left(x \mapsto \frac{f(x^{i\mapsto 1})-f(x^{i\mapsto -1})}{2}\right)$
- Influence of coordinate $i \in [n]$ on $f\colon \{1,-1\}^n \to \mathbb{R}$ is defined by

  $Inf_i(f) = \mathop{\mathrm{E}}_{x\sim\{1,-1\}^n}(D_i f(x)^2) = ||D_i f||_2^2$
- i is called relevant for $f\colon \{1,-1\}^n \to \mathbb{R}$ iff $Inf_i(f) > 0$
  i.e., $f\left(x^{i\mapsto 1}\right) \neq f\left(x^{i\mapsto -1}\right)$ for at least one $x \in \{1,-1\}^n$

Remark

- For Boolean f, $x \mapsto D_i f(x)^2$ is an indicator function for whether i is pivotal for f on x. So $Inf_i(f) = \mathop{\mathrm{E}}_{x\sim\{1,-1\}^n}(D_i f(x)^2)$
- This justifies the generalization above.

Derivative and Influence

Remark

Let $i \in [n]$ *and* $f : \{1,-1\}^n \to \mathbb{R}$. *Then*

1. $D_i f = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S) \chi_{S \setminus \{i\}}$
2. $Inf_i(f) = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S)^2$

Wake-Up-Question: Prove the remark

Derivative and Influence

Remark

Let $i \in [n]$ *and* $f: \{1,-1\}^n \to \mathbb{R}$. *Then*

1. $D_i f = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S) \chi_{S \setminus \{i\}}$
2. $Inf_i(f) = \sum_{\substack{S \subseteq [n] \\ i \in S}} \hat{f}(S)^2$

Wake-Up-Question: Prove the remark

Proof

1. Fourier-expand f in the definition of $D_i f$
2. Follows from 1.

Derivative and Influence

Proposition

Let $i \in [n]$ *and* $f: \{1,-1\}^n \rightarrow \{1,-1\}$ *. Then*

1. If f is monotone, then $Inf_i(f) = \hat{f}(\{i\})$
2. If additionally f is transitive-symmetric, then $Inf_i(f) \leq \frac{1}{\sqrt{n}}$

Proof

1. $Inf_i(f)$

- $= \Pr_{x\sim\{1,-1\}^n}\left(f(x) \neq f\left(x^{\oplus i}\right)\right)$
- $= \Pr_{x\sim\{1,-1\}^n}(f\left(x^{i\mapsto 1}\right) \neq f\left(x^{i\mapsto -1}\right))$
- $= \mathop{\mathrm{E}}_{x\sim\{1,-1\}^n}(\frac{f\left(x^{i\mapsto 1}\right)-f\left(x^{i\mapsto -1}\right)}{2})$ (due to monotony)
- $= E(D_i f)$
- $= \widehat{D_i f}(\emptyset) = \hat{f}(\{i\})$

2. $1 = \sum_{S\subseteq[n]} \hat{f}(S)^2$

- $\geq \sum_i^n \hat{f}(\{i\})^2$
- $= n\, \hat{f}(\{i\})^2$ (due to transitive symmetry)
- $= n\, Inf_i(f)^2$ (due to 1.)

ith Expectation and ith Laplacian

Definition

Let $i \in \{-1,1\}$. The ith Expectation E_i and Laplacian L_i on $\mathbb{R}^{\{1,-1\}^n}$ are defined by

- $E_i f(x) = E_{y\sim\{1,-1\}}(f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n))$
- $L_i f = f - E_i f$

Remark

Let $f: \{1,-1\}^n \to \mathbb{R}$ $and\, x \in \{1,-1\}^n$

- $E_i f(x) = \frac{f(x)+f(x^{\oplus i})}{2}$
- $f(x) = E_i f(x) + x_i D_i f(x) = E_i f(x) + L_i f(x)$
- $E_i f(x) = \sum_{\substack{S\subseteq[n] \\ i\notin S}} \hat{f}(S)\, x^S$
- $L_i f(x) = \sum_{\substack{S\subseteq[n] \\ i\notin S}} \hat{f}(S)\, x^S$
- $\langle L_i f, f\rangle = \langle L_i f, L_i f\rangle = Inf_i(f)$

Sensitivity

Definition

- The total influence of $f: \{1,-1\}^n \rightarrow \mathbb{R}$ is $Inf(f) = \sum_{i=}^{n} Inf_i(f)$
- The sensitivity $sens_f(x)$ of f: $\{1,-1\}^n \rightarrow \{1,-1\}$ at x is defined to be the number of pivotal coordinates for f on x

Theorem

Fix $n \in \mathbb{N}_0$. For $f: \{1,-1\}^n \rightarrow \{1,-1\}$

- $E_x(|\{i \in [n] \mid x_i = f(x)) = \frac{n}{2} + \frac{1}{2} \sum_{1 \le i \le n} \hat{f}(\{i\})$
- *and for odd n this maximized iff* $f = maj_n$, *hence among all monotone* $f: \{1,-1\}^n \rightarrow \{1,-1\}$ maj_n *is the one with maximal total influence.*

Proof :

- lhs = $\sum_{1 \le i \le n} \frac{1+E_x(f(x)x_i)}{2} = \frac{n}{2} + \frac{1}{2} \sum_{1 \le i \le n} \langle f, \chi_{\{i\}} \rangle$ = rhs
- $\frac{1}{2} \sum_{1 \le i \le n} \hat{f}(\{i\}) = E_x\big(f(x)(x_1 + \cdots + x_n)\big) \le E_x(|x_1 + \cdots + x_n|)$ where equality holds iff $f(x) = \text{sgn}(x_1 + \cdots x_n)$ for all $x \in \{1,-1\}^n$ with $(x_1 + \cdots x_n) \neq 0$. But if n is odd, then $(x_1 + \cdots x_n) \neq 0$ for all $x \in \{1,-1\}^n$

Getting famous with Fourier analysis research...

- A complexity measure related to sensitivity plays a prominent role in a recent breakthrough result described in a short paper (Huang 19)
- Concerns the sensitivity conjecture (Nisan/Szegedy 92)
  - Roughly: Most complexity measures on boolean functions could be shown to be polynomially reducible to each other
  - For a sensitivity based complexity this could not be proved - until Huang's insight
- There are many nice explanations on various theory-related blogs (see here for the links) and even a short twitter explanation by O'Donnell.

Discrete gradient and Laplacian

Definition

- The discrete gradient operator is defined by

$$\nabla\colon\ \mathbb{R}^{\{1,-1\}^n} \to (\mathbb{R}^n)^{\{1,-1\}^n}, f \mapsto \left(x \mapsto \begin{pmatrix} D_1 f(x) \\ \dots \\ D_2 f(x) \end{pmatrix} \right)$$

- The Laplacian is $L = \sum_i^n L_i$

Remark

Let $f\colon \{1,-1\}^n \to \mathbb{R}$

- $Lf = \sum_{S \subseteq [n]} |S| \hat{f}(S)\, \chi_S$
- $\langle Lf, f \rangle = I(f)$
- $I(f) = \sum_{S \subseteq [n]} |S| \hat{f}(S)^2 = \sum_{k=0}^n k \left|\left| f_{=k} \right|\right|_2^2$

Discrete gradient and Laplacian

Remark

If $f:\{1,-1\}^n \to \{1,-1\}$ *and* $x \in \{1,-1\}^n$

- $\|\nabla f(x)\|_2^2 = sens_f(x)$
- $Lf(x) = f(x)sens_f(x)$
- $I(f) = E_{S\sim \hat{f}^2}(|S|)$

Proposition (Poincare Lemma)

Let $f:\{1,-1\}^n \to \mathbb{R}$. *Then* $Var(f) \leq I(f)$

Proof :

- $Var(f) = \sum_{\substack{S\subseteq[n] \\ S\neq\emptyset}} \hat{f}(S)^2 \leq \sum_{S\subseteq[n]} |S|\hat{f}(S)^2 = I(f)$

NOISE STABILITY AND ARROW‘S PROOF OF THEOREM

Noise Stability: Motivation

- $f: \{1, -1\}^n \to \{1, -1\}$ as voting rule with 2 candidates and n voters
- Assume impartial culture assumption:
  votes $x = (x_1, \dots, x_n)$ chosen independently
- Now assume noise in misrecording y_i of vote x_i with chance $1 - \rho, \rho \in [0,1]$
- Want to know whether noise effects outcome, i.e., what is probability of $f(x) = f(y)$?
- Leads to notion of noise stability

Correlated sampling

Definition

- For $\rho \in [0,1]$ and fixed $x \in \{1,-1\}^n$ the sample $y \sim N_\rho(x)$ is drawn as follows :
  - $y_i = x_i$ with probability ρ
  - $y_i =$ uniformly random with probability $1-\rho$
- More generally for $\rho \in [-1,1]$ and fixed $x \in \{1,-1\}^n$ $y \sim N_\rho(x)$ is drawn as follows :
  - $y_i = x_i$ with probability $\frac{1}{2}+\frac{1}{2}\rho$
  - $y_i = -x_i$ with probability $\frac{1}{2}-\frac{1}{2}\rho$

We say that y is ρ-correlated to x

- If $x \sim \{1,-1\}^n$ and $y \sim N_\rho(x)$ then (x,y) is a ρ-correlated pair. In these slides we abbreviate this with $(x,y) \approx \rho$

This is equivalent to saying $E(x_i)=0, E(y_i)=0$ and $E(x_i,y_i)=\rho$ for each i

Noise Stability: Definition

$f: \{1,-1\}^n \to \{1,-1\}$ as voting rule with 2 candidates and n voters

Definition

For $f: \{1,-1\}^n \to \mathbb{R}$ and $\rho \in [-1,1]$ the noise stability of f at ρ is

$$Stab_\rho(f) = E_{(x,y)\approx\rho}(f(x)f(y))$$

Remark

If $f: \{1,-1\}^n \to \{1,-1\}$ we have

$$Stab_\rho(f) = \Pr_{(x,y)\approx\rho}(f(x) = f(y)) - \Pr_{(x,y)\approx\rho}(f(x) \neq f(y))$$

$$= 2\Pr_{(x,y)\approx\rho}\big(f(x) = f(y)\big) - 1$$

Example

- The constant functions have noise stability 1 at each $\rho \in [-1,1]$
- For dictators
  $Stab_\rho(\chi_i) = \rho$ for all $\rho \in [-1,1]$
- More generally for parities one has

$$Stab_\rho(\chi_S) = E_{(x,y)\approx\rho}(x^S y^S) = E_{(x,y)\approx\rho}\left(\prod_{i\in S} x_i y_i\right)$$

$= \prod_{i\in S} E_{x_i,y_i}(x_i y_i)$ (by (independece of (x_i, y_i) accross i)

$= \prod_{i\in S} \rho = \rho^{|S|}$

Noise operator

Definition

Let $\rho \in [-1,1]$. The noise operator T_ρ with parameter ρ is the vector space endomorphism of $\mathbb{R}^{\{1,-1\}^n}$ defined by

$T_\rho f(x) = E_{y\sim N_\rho(x)}(f(y))$

for all $f: \{1,-1\}^n \to \mathbb{R}$ and $x \in \{1,-1\}^n$

Theorem

For $\rho \in [-1,1]$ *and* $f: \{1,-1\}^n \to \mathbb{R}$ *:*

$$T_\rho f = \sum_{S\subseteq[n]} \rho^{|S|}\hat{f}(S)\chi_S = \sum_{k=0}^{n} \rho^k f_{=k}$$

Proof

- By linearity it is sufficient to prove $T_\rho\chi_S = \rho^{|S|}\chi_S$, but this follows from
- $T_\rho\chi_S(x) = E_{y\sim N_\rho(x)}(y^S) = \prod_{i\in S} E_{y\sim N_\rho(x)}(y_i) = \prod_{i\in S}(\rho x_i) = \rho^{|S|}\chi_S$
- Here we used that $E_{y\sim N_\rho(x)}(y_i) = \left(\frac{1}{2}+\frac{1}{2}\rho\right)x_i + \left(\frac{1}{2}-\frac{1}{2}\rho\right)(-x_i) = \rho x_i$

Stability and Noise operator

Theorem

For $\rho \in [-1,1]$ *and* $f: \{1,-1\}^n \rightarrow \mathbb{R}$

$$Stab_\rho(f) = \langle f, T_\rho f \rangle$$

Proof

- $Stab_\rho(f) = \mathrm{E}_{(x,y)\approx\rho}\big(f(x)f(y)\big) = E_{x\sim\{1,-1\}^n}(f(x)E_{y\sim N_\rho(x)}(f(y)))$

Corollary

For $\rho \in [-1,1]$ *and* $f: \{1,-1\}^n \rightarrow \mathbb{R}$

$Stab_\rho(f) = \sum_{S\subseteq[n]} \rho^{|S|} \hat{f}(S)^2 = \sum_{k=0}^{n} \rho^k \left|\left|f_{=k}\right|\right|_2^2$

In particular

$Stab_\rho(f) = E_{S\sim\hat{f}^2}\left(\rho^{|S|}\right)$ *for all* $f: \{1,-1\}^n \rightarrow \{1,-1\}$

Condorcet election

- For two candidates, majority function has all good properties
- For at least 3 candidates problem of social becomes much mor difficult
- Remember Condorcet election:
  - Compare each pair of alternatives
  - Declare “a” is socially preferred to “b” if more voters strictly prefer a to b
- Condorcet winner: Wins all of the pairwise elections in which he participates (for 3 candidates there are two such pairwise elections in which he participates).

Boolean Encoding of Condorcet

- Encode preference on candidates in a pairwise election by $\{1, -1\}$
- Encode a ranking of an individual voter w.r.t. set of candidates $\{A, B, C\}$ by a 3-tuple of consistent preferences, i.e., by an element of the set 6-element set

$$R = \{(\quad A\ (1)\ vs.\ B\ (-1)?\ , B(1)\ vs\ \ C(-1)?\ , C\ (1)\ vs.\ A\ (-1)\quad)\} =$$
$$= \{(1,1,-1), (1,-1,-1), (-1,1,-1), (-1,1,1), (1,-1,1), (-1,-1,1)\}$$

- E.g. $(1,1,-1)$ encodes ranking: $A < B < C$:
  A preferred to B, B preferred to C (and, consistently, A preferred to C).

Example

Three voters (n= 3), three candidates, f = maj_n with existing Condorcet winner a:

	Voter rankings				
	#1	#2	#3		Societal aggregation
A (1) vs. B (-1)	1	1	-1	= x	f(x) = 1
B (1) vs. C (-1)	1	-1	1	= y	f(y) = 1
C (1) vs. A (-1)	-1	-1	1	= z	f(z) = -1

Example

Three voters (n= 3), three candidates , f = maj_n **without** existing Condorcet winner

	Voter rankings				
	#1	#2	#3		Societal aggregation
A (1) vs. B (-1)	1	1	-1	= x	f(x) = 1
B (1) vs. C (-1)	1	-1	1	= y	f(y) = 1
C (1) vs. A (-1)	-1	1	1	= z	f(z) = 1

Societal outcome (1,1,1) not consistent (circular)

Theorem

Consider a 3-candidate Condorcet election using the same voting rule $f : \{-1,1\}^n \to \{-1,1\}$ for each pairwise election. If each of the n voters chooses uniformly and independently one of the 3! = 6 candidate rankings (of R), then the probability of a Condorcet winner is precisely : $\frac{3}{4} - \frac{3}{4}\, Stab_{-\frac{1}{3}} f$

Proof

- Let $x, y, z \in \{1,-1\}^n$ be the votes for the pairwise elections A vs B, B vs C, A vs. C
- By assumption (x_i, y_i, z_i) are chosen uniformly and independently out of R
- Function g: $\{-1,1\}^3 \to \{0,1\}$, $w \mapsto \frac{3}{4} - \frac{1}{4} w_1 w_2 - \frac{1}{4} w_1 w_3 - \frac{1}{4} w_2 w_3$ is indicator function for R
- Probability of Condorcet winner is

$$E[g(f(x), f(y), f(z))]$$
$$= \frac{3}{4} - \frac{1}{4} E[f(x)f(y)] - \frac{1}{4} E[f(x)f(z)] - \frac{1}{4} E[f(y)f(z)].$$

Theorem

Consider a 3-candidate Condorcet election using the same voting rule $f : \{-1,1\}^n \rightarrow \{-1,1\}$ for each pairwise election.
If each of the n voters chooses uniformly and independently one of the $3! = 6$ candidate rankings (of R), then the probability of a Condorcet winner is precisely : $\frac{3}{4} - \frac{3}{4}\, Stab_{-\frac{1}{3}}\, f$

Proof (continued)

- Probability of Condorcet winner is
$$E[g(f(x), f(y), f(z))] = \frac{3}{4} - \frac{1}{4}\, E[f(x)f(y)] - \frac{1}{4} E[f(x)f(z)] - \frac{1}{4} E[f(y)f(z)].$$
- Now $E[x_i] = 0 = E[y_i]$ and $E[x_i y_i] = \frac{2}{6} - \frac{4}{6} = -\frac{1}{3}$ for each i
- So x, y are $-\frac{1}{3}$ correlated, so $E\big(f(x)f(y)\big) = Stab_{-\frac{1}{3}}(f)$.
- Similarly $E\big(f(x)f(z)\big) = E\big(f(y)f(z)\big) = Stab_{-\frac{1}{3}}(f)$.

Theorem

In a 3-candidate Condorcet election using $f : \{-1,1\}^n \rightarrow \{-1,1\}$, the probability of a Condorcet winner *is at most:* $\frac{7}{9} + \frac{2}{9}\,||f_{=1}||_2^2$

Proof

- $\frac{3}{4} - \frac{3}{4}\,Stab_{-\frac{1}{3}} f$

  $= \frac{3}{4} - \frac{3}{4}\,(||f_{=0}||_2^2 - \frac{1}{3}\,(||f_{=1}||_2^2 + \frac{1}{9}\,||f_{=2}||_2^2 - \frac{1}{27}\,||f_{=3}||_2^2 + \cdots)$
- $\leq \frac{3}{4}\,(1 + \frac{1}{3}\,(||f_{=1}||_2^2 + \frac{1}{27}\,||f_{=3}||_2^2 + \frac{1}{243}\,||f_{=5}||_2^2 + \cdots)$
- $\leq \frac{3}{4}\,(1 + \frac{1}{3}\,(||f_{=1}||_2^2 + \frac{1}{27}(\,||f_{=3}||_2^2 + \,||f_{=5}||_2^2 + \cdots))$
- $\leq \frac{3}{4}\,(1 + \frac{1}{3}\,(||f_{=1}||_2^2 + \frac{1}{27}\left(1 - ||f_{=1}||_2^2\right) = \frac{7}{9} + \frac{2}{9}\,||f_{=1}||_2^2$

Theorem (Arrows' theorem)

Suppose $f : \{-1,1\}^n \to \{-1,1\}$ is a unanimous voting rule used in a 3-candidate Condorcet election. If there is always *a Condorcet winner, then f must be a dictatorship.*

Proof

- If there is always a Condorcet winner, then

$$1 \leq \frac{7}{9} + \frac{2}{9}\,||f_{=1}||_2^2 \leq \frac{7}{9} + \frac{2}{9}\,||f||_2^2 = \frac{7}{9} + \frac{2}{9} = 1$$

- Hence $||f_{=1}||_2^2 = 1 = ||f||_2^2$
- Hence $f = f_{=1}$
- But this can hold only if f is either a dictator or a negated dictator. (*)
- As f is unanimous, f is a dictator.

Wake-up Question

- Show (*): If $f = f_{=1}$ then f is either a dictator or a negated dictator.

Answer to Wake-up Question

- Show (*): If $f = f_{=1}$ then f is either a dictator or a negated dictator.
- Proof:
  - $f = \sum_{i\in[n]} \hat{f}(\{i\})x_i$
  - Feach $x \in \{1,-1\}^n$ and $i \in [n]$
    - Either $f(x) = f(x^{\oplus i})$.
    - Or $f(x) \neq f\left(x^{\oplus i}\right)$
  - In the first case $\left|\hat{f}(\{i\})\right| = 0$
  - In the second case $\left|\hat{f}(\{i\})\right| = 1$
  - As $\left|\left|f_{=1}\right|\right|_2^2 = 1$, exactly for one $i \in [n]$: $\left|\hat{f}(\{i\})\right| = 1$
  - So $f = \chi_i$ or $f = -\chi_i$ for some i

Uhhh, a lecture with a hopefully useful

APPENDIX

References

- (Kalai 02)
  G. Kalai. A fourier-theoretic perspective on the condorcet paradox and arrow's theorem. Advances in Applied Mathematics, 29(3):412–426, 2002.
- (Nalon/Spencer 04)
  N. Alon and J. Spencer. The Probabilistic Method. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 2004.
- (Huang 19)
  H. Huang. Induced subgraphs of hypercubes and a proof of the Sensitivity Conjecture. arXiv e-prints, page arXiv:1907.00847, July 2019.
- (Nisan/Szegedy 92)
  N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. In Proceedings of the Twenty-Fourth Annual ACM Symposium on Theory of Computing, STOC '92, pages 462–467, New York, NY, USA, 1992. Association for Computing Machinery.

Color Convention in this course

- Formulae, when occurring inline
- Newly introduced terminology and definitions
- Important results (observations, theorems) as well as emphasizing some aspects
- Examples are given with standard orange with possibly light orange frame
- Comments and notes
- Algorithms